UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/693,403	10/24/2003	James R. Stinger	10992509-2	1929

7590 04/30/2007

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P. O. Box 272400
Fort Collins, CO 80527-2400

EXAMINER	
COLAN, GIOVANNA B	
ART UNIT	PAPER NUMBER
2162	
MAIL DATE	DELIVERY MODE
04/30/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/693,403	STINGER, JAMES R.	
	Examiner	Art Unit	
	Giovanna Colan	2162	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *22 January 2007*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,3,5,7,9,12 and 15-20* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 3,5,7,9,12,15-20* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This action is issued in response to applicant filed request for continued examination (RCE) on 01/22/2007.

2. Claims 1, 7, and 12, have been amended. No claims were added. Claims 2, 4, 6, 8, 10 – 11, and 13 – 14 were canceled.

3. Claims 1, 3, 5, 7, 9, 12, and 15 – 20 are pending in this application.

4. Applicant's arguments with respect to amended claims 1, 7, and 12 have been considered but are moot in view of the new ground(s) of rejection.

Continued Examination Under 37 CFR 1.114

5. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 10/11/2006 has been entered.

Claim Rejections - 35 USC § 102

6. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

7. Claims 1, 3, 5, 7, 9, 12, and 15 – 20 are rejected under 35 U.S.C. 102(e) as being anticipated by Alam et al. (Alam hereinafter) (US Patent No. 6,336,124 B1, filed: October 1, 1998).

Regarding Claim 1 and 7, Alam discloses a computer-readable medium having stored thereon sequences of instructions, said sequences of instructions including instructions which, when executed by a processor, cause said processor to perform the steps of:

receiving a page description language representation of the document (Col. 6, lines 24 – 28, Alam) for providing a list of words in the document (Col. 6, lines 57 – 59, words, Alam) and position information for the words (Col. 6 and 7, lines 54 – 57 and 60 – 63; respectively, X and Y coordinates, Alam); and

automatically identifying table data in the document based on the page description language representation of the document and at least one table identifying feature (Col. 7, lines 26 – 28 and 32 – 35, Alam), wherein the identifying step includes:

dividing the document into one or more pages (Col. 15, lines 56 – 58, dividing into sub-pages, Alam);

dividing each page into a plurality of lines (Col. 15, lines 57 – 60, Alam);

for each line (Col. 8, lines 57 – 60, Alam), clustering the words of the line into one or more word clusters (Col. 9, lines 1 – 6, Alam[1]), wherein each cluster includes one or more words (Col. 16, lines 15 – 22; if the **block is a paragraph that ends with an incomplete sentence** or an improper termination ... whether **the block contains one or more sentences**, Alam), each cluster having a horizontal beginning point, horizontal midpoint, and horizontal end point (Col. 11, lines 8 – 17, wherein the right most X coordinate corresponds to the horizontal beginning point claimed; wherein the center X coordinate corresponds to the horizontal midpoint claimed; and wherein the left X coordinate corresponds to the horizontal end point claimed; Alam);

for clusters in the plurality of lines, comparing alignment of the horizontal beginning point, horizontal midpoint, and horizontal end point of clusters between lines, wherein a cluster in a first line is aligned with a cluster in a previous line if at least one of the horizontal beginning point, horizontal midpoint, and horizontal end point of the cluster in the first line is aligned with at least one of the horizontal

beginning point, horizontal midpoint, and horizontal end point of the cluster in the previous line (Col. 10, and 11, lines 24 – 33, and 4 – 7, "...To determine whether the selected line is near the current paragraph in the Y direction, the appropriate Y coordinate (s) of the selected line are compared with the appropriate Y coordinate(s) of the previous line of the current paragraph to determine whether certain parameters and/or thresholds are satisfied..."; and "... comparing the left X coordinate of the first word of the current line with the left X coordinate of the first word of a first line in the current paragraph to account for the hanging indent..."; respectively, Alam); and

identifying a line as being part of a table in response to more than one cluster of the line being aligned with clusters of previous lines identified as part of the table (Col. 12, and 17, 47 – 49 and 51 – 52, and 10 – 18; "...determines if the current block is not a table, step 1904 breaks up the current block into elements such that ach element can be displayed within the display configuration. Each element of a paragraph block may be, for example, a word contained in the paragraph. Other division of a block into elements may be implemented. For example, each element of a list block may be an item or a line in the list..."; respectively, Alam); and

outputting data descriptive of the lines of the table (Col. 1 – 2, and 19, lines 65 – 67 and 1, and 48 – 60; "...at least a portion of the first row forming the row heading of sample table 2200 is displayed in each of the display pages...", Alam).

---

[1] Wherein examiner interprets the step of determining the spacing between words as the step of

Regarding Claim 3, Alam discloses a method wherein the step of automatically identifying table data in the document based on the number of word clusters for each line and the alignment of the word clusters between lines (Col. 12, lines 47 – 49 and 51 – 52, Alam) further comprises:

using the word clusters to generate column position information (Col. 9 and 18, lines 10 – 12 and 9 – 12; respectively, Alam), wherein the column information includes for each column a horizontal beginning point, horizontal midpoint, and horizontal end point (Col. 11, lines 8 – 17, wherein the right most X coordinate corresponds to the horizontal beginning point claimed; wherein the center X coordinate corresponds to the horizontal midpoint claimed; and wherein the left X coordinate corresponds to the horizontal end point claimed; Alam); and

updating the column position information (Col. 10, lines 42 – 45, Alam[2]) by performing a union operation between the column position information of a previous line (Col. 10, lines 26 – 29, Y coordinates of the previous lines, Alam) and the column position information of a current line (Col. 10, lines 24 – 27, Y coordinates of the selected lines, Alam).

Regarding Claim 5, Alam discloses a method wherein receiving a page description language representation of the document for providing a list of words in the document and position information for the words includes receiving a PDF

---

clustering the words claimed.

representation of the document (Col. 2, lines 1 – 5, portable document format (PDF), Alam), and wherein converting the table data encompassed by each table bounding box to a markup language representation includes converting the table data encompassed by each table bounding box to a HTML representation (Col. 6, lines 42 – 47, Alam).

Regarding Claim 9, Alam discloses a computer-readable medium further containing instructions which, when executed by said processor, would cause said processor to perform the steps of:

using the word clusters to generate column position information (Col. 9 and 18, lines 10 – 12 and 9 – 12; respectively, Alam), wherein the column information includes for each column a horizontal beginning point, horizontal midpoint, and horizontal end point (Col. 11, lines 8 – 17, wherein the right most X coordinate corresponds to the horizontal beginning point claimed; wherein the center X coordinate corresponds to the horizontal midpoint claimed; and wherein the left X coordinate corresponds to the horizontal end point claimed; Alam); and

updating the column position information (Col. 10, lines 42 – 45, Alam[3]) by performing a union operation between the column position information of a previous line (Col. 10, lines 26 – 29, Y coordinates of the previous lines, Alam) and the column position information of a current line (Col. 10, lines 24 – 27, Y coordinates of the selected lines, Alam).

---

[2] Wherein examiner interprets the step of determining that the line is not near the current paragraph as the step of updating position information claimed.
[3] Wherein examiner interprets the step of determining that the line is not near the current paragraph as the step of updating position information claimed.

Regarding Claim 12, Alam discloses a document processing system comprising:

a processor for executing programs (Fig. 2, item 151, Col. 4, lines 61 – 64, Alam); and

a table identification program (Col. 4, lines 43 – 46, Alam) for receiving a page description language representation of a document (Col. 6, lines 24 – 28, Alam), the page description language representation providing a list of words in the document (Col. 6, lines 57 – 59, words, Alam) and position information for the words (Col. 6 and 7, lines 54 – 57 and 60 – 63; respectively, X and Y coordinates, Alam), and for automatically identifying table data in the document based on the page description language representation of the document and at least one table identifying feature (Col. 7, lines 26 – 28 and 32 – 35, Alam), wherein the identification program is configured to,

dividing the document into one or more pages (Col. 15, lines 56 – 58, dividing into sub-pages, Alam);

dividing each page into a plurality of lines (Col. 15, lines 57 – 60, Alam);

for each line (Col. 8, lines 57 – 60, Alam), cluster the words of the line into one or more word clusters (Col. 9, lines 1 – 6, Alam[4]), wherein each cluster includes one or more words (Col. 16, lines 15 – 22; if the **block is a paragraph that ends with an incomplete sentence** or an improper termination ... whether **the block contains one or more sentences**, Alam), each cluster having a horizontal beginning point, horizontal midpoint, and horizontal end point (Col. 11,

lines 8 – 17, wherein the right most X coordinate corresponds to the horizontal beginning point claimed; wherein the center X coordinate corresponds to the horizontal midpoint claimed; and wherein the left X coordinate corresponds to the horizontal end point claimed; Alam);

for clusters in the plurality of lines, compare alignment of the horizontal beginning point, horizontal midpoint, and horizontal end point of clusters between lines, wherein a cluster in a first line is aligned with a cluster in a previous line if at least one of the horizontal beginning point, horizontal midpoint, and horizontal end point of the cluster in the first line is aligned with at least one of the horizontal beginning point, horizontal midpoint, and horizontal end point of the cluster in the previous line (Col. 10, and 11, lines 24 – 33, and 4 – 7, "...To determine whether the selected line is near the current paragraph in the Y direction, the appropriate Y coordinate (s) of the selected line are compared with the appropriate Y coordinate(s) of the previous line of the current paragraph to determine whether certain parameters and/or thresholds are satisfied..."; and "... comparing the left X coordinate of the first word of the current line with the left X coordinate of the first word of a first line in the current paragraph to account for the hanging indent..."; respectively, Alam); and

identify a line as being part of a table in response to more than one cluster of the line being aligned with clusters of previous lines identified as part of the table (Col. 12, and 17, 47 – 49 and 51 – 52, and 10 – 18; "...determines if the

[4] Wherein examiner interprets the step of determining the spacing between words as the step of

current block is not a table, step 1904 breaks up the current block into elements such that ach element can be displayed within the display configuration. Each element of a paragraph block may be, for example, a word contained in the paragraph. Other division of a block into elements may be implemented. For example, each element of a list block may be an item or a line in the list..."; respectively, Alam); and

output data descriptive of the lines of the table (Col. 1 – 2, and 19, lines 65 – 67 and 1, and 48 – 60; "...at least a portion of the first row forming the row heading of sample table 2200 is displayed in each of the display pages...", Alam).

Regarding Claim 15, Alam discloses a document processing system of claim 13 wherein the table identification program further comprises:

a conversion module (Fig. 6, item 628, Col. 6, lines 41 – 45, converter, Alam) coupled to the bounding box generation module for receiving the table bounding box for each table in the document (Fig. 6, item 612, Col. 6, lines 35 – 40, ACROBAT CAPTURE and ACROBAT WRITER, Alam), and for converting the words encompassed by the table bounding box into a markup language representation that maintains the table structure of each table (Col. 6, lines 42 – 47, Alam).

Regarding Claim 16, Alam discloses a method wherein the step of automatically identifying table data in the document based on the page description language

---

clustering the words claimed.

representation of the document and at least one table identifying feature further comprises:

automatically identifying table data in the document based on one or more table headings (Col. 2, lines 39 – 44, Alam).

Regarding Claim 17, Alam discloses a method wherein the step of automatically identifying table data in the document based on the page description language representation of the document and at least one table identifying feature further comprises:

automatically identifying table data in the document based on one or more horizontal lines (Fig 15B, Col. 14, lines 38 – 43, five rows, Alam) and vertical lines that separate rows or columns of the table (Fig. 15C, Col. 14, lines 46 – 49, two columns, Alam).

Regarding Claim 18, Alam discloses a method, wherein the step of automatically identifying table data in the document based on the number of word clusters for each line and the alignment of the word clusters comprises:

determining whether the number of word clusters in a line is greater than a threshold value (Col. 8, lines 21 – 26; If the inter-word spacing or distance in the Y direction is greater than a threshold…; Alam); and

classifying the word clusters in the line as a row of a table in response to the number of word clusters in a line being greater than the threshold value (Col. 8, lines 28 – 34; Alam).

Regarding Claim 19, Alam discloses a computer-readable medium, wherein the instructions for automatically identifying table data in the document based on the number of word clusters for each line and the alignment of the word clusters include instructions that when executed by a processor cause the processor to perform the steps further comprising:

determining whether the number of word clusters in a line is greater than a threshold value (Col. 8, lines 21 – 26; If the inter-word spacing or distance in the Y direction is greater than a threshold…; Alam); and

classifying the word clusters in the line as a row of a table in response to the number of word clusters in a line being greater than the threshold value (Col. 8, lines 28 – 34; Alam).

Regarding Claim 20, Alam discloses a document processing system, wherein the table identification program is further configured to:

determine whether the number of word clusters in a line is greater than a threshold value (Col. 8, lines 21 – 26; If the inter-word spacing or distance in the Y direction is greater than a threshold…; Alam); and

classify the word clusters in the line as a row of a table in response to the number of word clusters in a line being greater than the threshold value (Col. 8, lines 28 – 34; Alam).

Prior Art Made Of Record

1. Alam et al. (US Patent No. 6,336,124 B1, filed: October 1, 1998) discloses conversion data representing a document to other formats for manipulation and display.

2. Egger et al. (US Patent No. 6,233,571 B1) discloses a method and apparatus for indexing searching and displaying data.

3. Hassan Alam (US Patent No. 5,737,442, issued: April 1998) discloses a processor based method for extracting tables from printed documents.

Points Of Contact

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Giovanna Colan whose telephone number is (571) 272-2752. The examiner can normally be reached on 8:30 am - 5:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Giovanna Colan
Examiner
Art Unit 2162
April 24, 2007

SANA AL-HASHEMI
PRIMARY EXAMINER